Sandia National Laboratories

Fact Sheet

Distributed, Intelligent RAS System for Large Computational Clusters

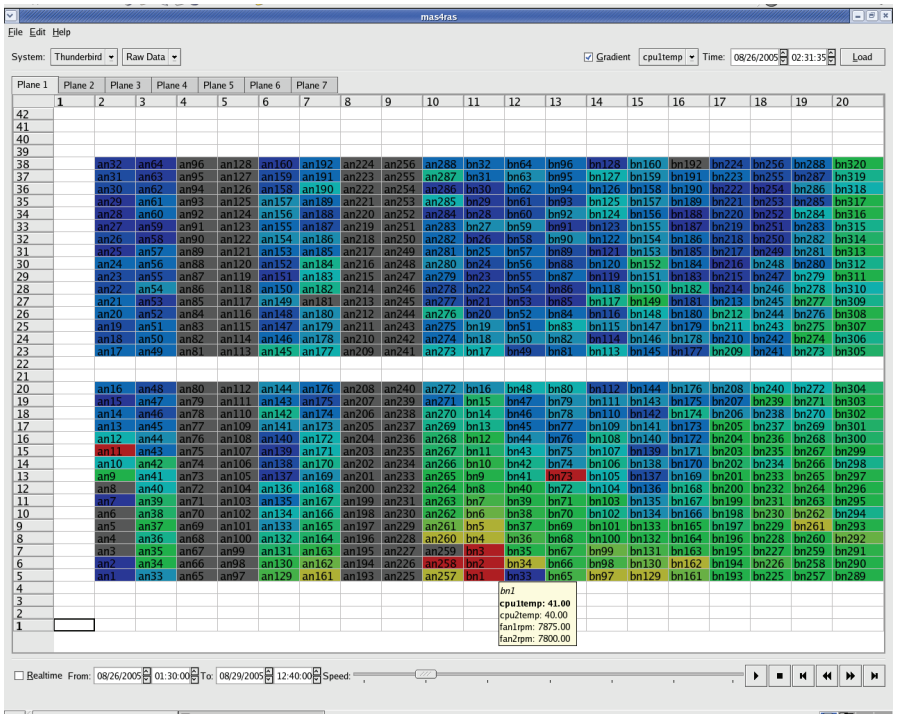


Figure 1. OVIS display showing run-time conditions and configuration of the Thunderbird cluster. Nodes are shown in the physical layout of the cluster racks. Values of raw and derived quantities are displayed by color coding of the nodes. Patterns and outliers are easily spotted by the eye.

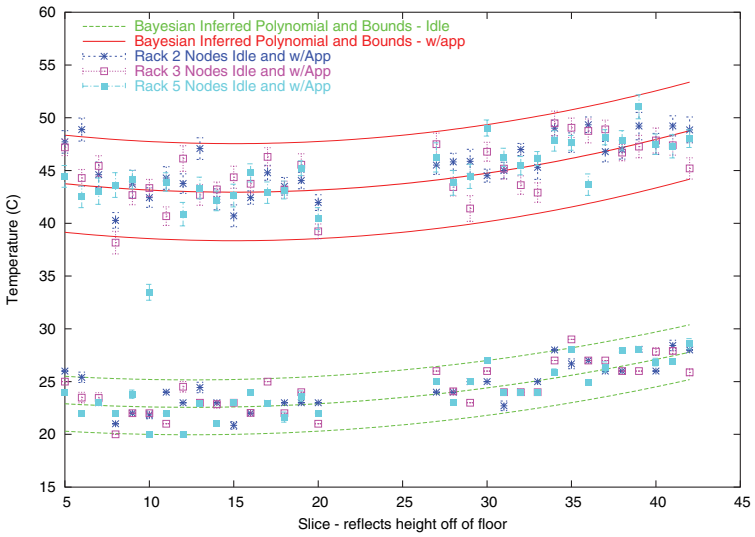
Cluster computing as the backbone of Sandia's capacity computing has become crucial to many of Sandia's missions. Viewing a cluster as a large collection of statistically similar devices allows us to detect aberrant node behavior due to various effects long before catastrophic failure occurs. This view is the basis for our Distributed, Intelligent RAS System for Large Computational Clusters project. RAS stands for Reliability, Availability, and Serviceability, and refers to the usability and stability of clusters. Our software tool, OVIS, allows a system administrator to make sense of environmental data even in the absence of a fundamental understanding of the total internal state of the nodes, the interaction between nodes, and the interaction of the machines with their environment. In addition, to advance problem detection, OVIS allows visualization of various configuration effects and can aid in their resolution.

OVIS performs statistical calculations on system and environmental data, characterizing single node behaviors based upon the behaviors of the entire set of nodes. Abnormal behavior is then automatically determined by detecting behaviors that have low statistical probabilities. An extremely simple example of this is the detection of a node's temperature value that is low compared to the temperature values of all other nodes in the system, when all nodes are situated in a uniform environment and under the same computational load. Here low can be defined as being a certain number of standard deviations away from the mean value of the all nodes' temperatures. This is in contrast to traditional methods that merely check for the crossing of some threshold value. Our statistical methods can detect problems earlier than the traditional methods.

SAND2006-2547P
April 2006

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories



OVIS determines representative models for system data from which abnormalities can be automatically detected. Here, raw data and the resulting models are shown for CPU temperature dependence on height in the Shasta Cluster. The models take the form of T ~ N(Q(h),σ), where h and T respectively denote height and temperature and N(Q(h),σ) is the normal distribution with mean Q(h) and variance sigma. In this case, Q is a quadratic polynomial. The coefficients of Q and the value of σ are identified by Bayesian inference. A problem with the node in slice 10, rack 5 is automatically identified as its value is outside the 95% probability bounds given the inferred model.

OVIS additionally employs a novel Bayesian inference mechanism that can create representative models of the system data. The models include both the most likely description of the data and the probability distribution around it. An example of this is a model consisting of of a normal distribution of temperatures about a mean, where the mean is a polynomial depending on height. Abnormalities can then be determined by comparisons to the representative model and by consideration of the probability distribution inferred. These models can also be used to isolate non-uniform external environmental parameters from the abnormalities within the nodes themselves, thus allowing additional analysis by simpler statistical methods.

The OVIS visual display is designed to capitalize on a humans' ability to efficiently spot patterns and abnormalities. The display shows the actual spatial configuration of the cluster. Encoding data values as colors which are mapped onto node positions on the display gives the user a very intuitive view of the data and how it relates to geography and application state.

It is a quick-and-easy, non-computationally intensive way to gain the benefits of considering the cluster as a comparative ensemble, rather than as singleton nodes. The color mappings are user-defined but, in general, low values are mapped to bluer colors and higher values move toward the red end of the spectrum. There are options to map binned values to particular colors to allow easy discrimination between values falling near boundaries (e.g. two σ from a mean boundary) or to produce a gradient display which facilitates better understanding of the distribution of values especially where spatial and temporal gradients are present. The gradient view can be put to good use in optimizing cold air distribution and choosing node groups and axis for application of Bayesian modeling techniques.

OVIS is intended to be generic enough to be used on any computational platform that provides enough nodes such that statistical methods are valid. To this end, OVIS has a configuration tool that allows the user to easily build the display replica of the cluster. All mainstream data collection methods will be supported (Ganglia, IPMI, Supermon, in band daemon collecting from /proc).

In our current usage, we gather and process data that is related to thermal parameters as temperature is an important factor affecting longevity and an indicator of several modes of failure in compute nodes. We will be addressing other parameters such as memory error rates, voltages, and network errors using the same techniques.

To date we have run OVIS as a single node collector/analyzer on Sandia platforms such as ICC/NWCC and Thunderbird. Due to scalability issues we are also developing a parallel version which will not only address these issues but also eliminate the single point of failure problem. Longer term, we will incorporate Bayesian networks to aid in diagnosis of problems.

For more information, contact
Jim Brandt (925) 294-2348
Ann Gentile (925) 294-3614
Phillipe Pébay (925) 294-2024
Matthew Wong (925) 294-3717